Research Report 1422

The Unit-Conduct of Fire Trainer (U-COFT) as a Medium for Assessing Gunner Proficiency: Test Reliability and Utility

Scott E. Graham

ARI Field Unit at Fort Knox, Kentucky

Training Research Laboratory

U. S. Army

Research Institute for the Behavioral and Social Sciences

March 1986

U. S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

WM. DARRYL HENDERSON
COL, IN
Commanding

Technical review by

John A. Boldovici
John C. Morey

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-POT, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ARI Research Report 1422	2. GOVT ACCESSION NO.: A169196
3. RECIPIENT'S CATALOG NUMBER:	
4. TITLE (and Subtitle): THE UNIT-CONDUCT OF FIRE TRAINER (U-COFT) AS A MEDIUM FOR ASSESSING GUNNER PROFICIENCY: TEST RELIABILITY AND UTILITY	5. TYPE OF REPORT & PERIOD COVERED: Final Report June 1985 - December 1985
	6. PERFORMING ORG. REPORT NUMBER: --
7. AUTHOR(s): Scott E. Graham	8. CONTRACT OR GRANT NUMBER(s): --
9. PERFORMING ORGANIZATION NAME AND ADDRESS: U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 2Q263744A795, 351 H2
11. CONTROLLING OFFICE NAME AND ADDRESS: U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600	12. REPORT DATE: March 1986 13. NUMBER OF PAGES: 21
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): --	15. SECURITY CLASS. (of this report): Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: --

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

--

18. SUPPLEMENTARY NOTES

--

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Armor	Simulation	Test construction
Army	Simulators	
Evaluation	Tanks	
Performance tests	Test and evaluation	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A device-mediated M1 tank gunnery test was developed for administration on the Unit-Conduct of Fire Trainer (U-COFT). The purpose of the research was to examine the reliability of the test, and to assess the utility of the test for estimating gunners' proficiency independently of the Tank Commander's (TC) contribution. The reliability of six of nine U-COFT measures exceeded .70.

The Gunners' hit rates were found to be heavily influenced by the TCs' performance, including the TCs' ability to train. The results show (Continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

ARI Research Report 1422

20. (Continued)

that TC performance must be stabilized if the U-COFT tests are to be used to assess Gunners' performance alone. Recommendations were made as to how the U-COFT could be efficiently used in units as a training and testing device.

Research Report 1422

The Unit-Conduct of Fire Trainer (U-COFT) as a Medium for Assessing Gunner Proficiency: Test Reliability and Utility

Scott E. Graham

ARI Field Unit at Fort Knox, Kentucky
Donald F. Haggard, Chief

Training Research Laboratory
Seward Smith, Acting Director

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

March 1986

Army Project Number
2Q263744A795

Training and Simulation

ARI Research Reports and Technical Reports are intended for sponsors of R&D tasks and for other research and military agencies. Any findings ready for implementation at the time of publication are presented in the last part of the Brief. Upon completion of a major phase of the task, formal recommendations for official action normally are conveyed to appropriate military agencies by briefing or Disposition Form.

FOREWORD

The U.S. Army Research Institute (ARI) Fort Knox Field Unit performs research and development to improve armor training and evaluation. Because tank gunnery skills are extremely critical and have high associated training costs, considerable emphasis has been placed on the development of electronic training devices and simulators. Within armor, the focus has been on the high-fidelity Unit-Conduct of Fire Trainer (U-COFT).

This report demonstrates how the U-COFT can be used as a device for evaluating gunnery performance as well as for training the requisite knowledge and skills. The psychometric properties of these device-mediated tests are found to compare favorably with other gunnery performance tests. In addition, the expanded testing applications of the U-COFT which are described potentially increase the simulator's cost-effectiveness.

EDGAR M. JOHNSON
Technical Director

THE UNIT-CONDUCT OF FIRE TRAINER (U-COFT) AS A MEDIUM FOR ASSESSING GUNNER PROFICIENCY: TEST RELIABILITY AND UTILITY

EXECUTIVE SUMMARY

Requirement:

The development of electronic training devices and simulators such as the Unit-Conduct of Fire Trainer (U-COFT) provides new test media for assessing M1 tank skills and knowledge. The objectives of this research were to (a) develop and use a U-COFT mediated tank gunnery test, (b) examine the reliability of the test, and (c) assess the utility of the test for estimating gunners' proficiency independently of the Tank Commanders' (TC) contributions.

Procedure:

A U-COFT Gunner's test was developed using a matrix sampling approach to match engagement conditions found in Table VIII of the M1 tank combat tables. The test and retest, consisting of 31 engagements in different sequences, were fired by inexperienced gunners who were paired with one of the three TCs. Test-retest reliability coefficients were computed for the nine U-COFT performance measures which included hit rate and target identification time.

Findings:

The reliability of six of the U-COFT measures exceeded .70. Tests using the U-COFT were therefore found to be a potentially valid means for assessing gunnery performance. The advantages of device-mediated tests were also discussed along with additional uses of U-COFT tests.

The gunners' hit rates were found to be heavily influenced by the TCs' performance, including the TCs' ability to train. Recommendations were made as to how the U-COFT could be efficiently used in units as a training and testing device.

Utilization of Findings:

The results of this research demonstrating the desirable psychometric qualities of the U-COFT as a testing device have been given to the Armor School and are being used as the foundation of other U-COFT testing projects. These projects include an evaluation of an Excellence Track training program in the 1st Armor Training Brigade and the validation of Project A psychomotor tests for predicting gunnery performance of Armor Officer Basic soldiers.

THE UNIT-CONDUCT OF FIRE TRAINER (U-COFT) AS A MEDIUM FOR ASSESSING GUNNER PROFICIENCY: TEST RELIABILITY AND UTILITY

CONTENTS

LIST OF TABLES

LIST OF FIGURES

THE UNIT-CONDUCT OF FIRE TRAINER (U-COFT) AS A MEDIUM FOR ASSESSING GUNNER PROFICIENCY: TEST RELIABILITY AND UTILITY

The Armor community is striving to improve its selection and training of M1 Tank Commanders (TC) and gunners. If these goals are to be accomplished, valid performance measures must be established which assess the full range of tank crewmen duties. The proliferation of electronic training devices and simulators provides new test media for measuring many of the skills and knowledge required for good tanker performance. It is appropriate, therefore, if not necessary, that the psychometric characteristics of these media be examined before they are implemented.

The characteristics which make a training device a good test medium are not the same as the characteristics which make a device good for training. The minimal requirement for a good training device is that practice with the device yields positive transfer to its "parent" weapon system. By contrast, the minimal requirement of an acceptable test medium is that it yields scores which are both reliable and valid.

Reliability refers to the consistency or dependability of a measure. A test cannot yield much value if scores fluctuate as a function of extraneous variables, such as differing test conditions or poor scoring. Reliability is important in testing for the same reasons it is important in other measurement operations, for example, measuring length. No confidence could be placed in a ruler which does not consistently distinguish between long boards and short boards. Likewise, little confidence can be placed in a test which does not consistently discriminate between superior and inferior performance.

Reliability is also important in that it places a limit on the validity of a test. Validity generally refers to questions of whether a test is actually measuring what it purports to be measuring. In psychometric terms, validity reflects the accuracy of estimating values of a population from which the test is a sample. The validity, or estimate of the population, cannot exceed a test's reliability. A test with high reliability is therefore potentially more valid than a test with low reliability. A reliable test is, however, not necessarily more valid as it might be measuring skills which are irrelevant to the test's intent, such as predicting criterion performance. A device-mediated tank gunnery test could, for example, very reliably measure knowledge or skills which are necessary for performance on the device, but which are unimportant to actual performance on the tank.

Two factors which affect the reliability of a test are standardization and test length. Tests will be more reliable when testing conditions are held as constant as possible, and when procedures used for scoring are consistent across test items and examinees. Also, the longer the test, that is, the more items included on the test, the more reliable it will be. Modern training devices therefore seem well suited as media for reliable tests. Standardized administration and scoring characterize nearly all of these devices, and increases in test length can be achieved inexpensively. Boldovici and Sabat (1985) provide additional discussions of reliability and device-mediated testing and training.

The Unit-Conduct of Fire Trainer (U-COFT) has recently been developed by the Simulation and Control Systems Division of General Electric Company in a major effort to augment armor training. The cost-effectiveness of the simulator will be enhanced if the U-COFT can be used for testing as well as for training. The U-COFT system already includes tests which makes recommendations for training remediation and rapid advancement. Additional plans are being made to use the U-COFT as a test medium in other capacities; such as estimating tank gunnery proficiency, and the validation of other less expensive armor performance tests.

Another desirable characteristic of tests which use the U-COFT is the utility of such tests for separating the contributions of individual crewmen in tank gunnery engagements. Assessing the relative contributions of tank commanders (TCs) and gunners is difficult in tank table exercises which yield collective measures, such as proportion of hits and time to engage. Such measures, like final scores of football games, are valuable for estimating team proficiency. They are not, however, particularly useful for identifying performance strengths and weaknesses of individual crewmen, which is necessary for efficient advancement, remediation, and personnel allocation.

Rationale and Purpose

Because the U-COFT holds considerable potential as a medium for assessing the proficiency of armor crewmen, research was undertaken to develop and use U-COFT mediated tank gunnery tests in ways which would permit:

1. Examining the reliability of the tests.

2. Assessing the utility of the tests for estimating gunners' proficiency independently of TCs' contributions.

METHOD

Participants

Thirty-two M60A3 loaders and drivers from the 194th Armor Brigade at Ft Knox, KY served as subjects. Most had ranks of Private First Class, and with few exceptions, had not served as gunners other than in Advanced Individual Training. None had experience on the M1 tank.

Four persons served as U-COFT TCs. Three were predominately used, with a fourth being used for one session. One civilian TC was an ex-General Electric employee who had hundreds of hours of U-COFT experience as a U-COFT TC (COFT-experienced). The other two were a Sergeant First Class from the Armor School Weapons Department (Sr NCO) and a Sergeant from the 2/6 Cavalry, a training support unit of the Armor School (Jr NCO). They had no prior U-COFT experience. In addition, three U-COFT Instructor/Operators (I/O) assisted in the research.

Materials and Apparatus

The M1 U-COFT is a high-fidelity whole-task gunnery trainer which presents computer-generated target imagery for training in normal and degraded operational modes. Training is directed by an Instructional Subsystem which includes a library of preprogrammed exercises and an adaptive Evaluation System for evaluating crew progress. The system centers around two training matrices, including a Tank Commander/Gunner matrix which contains 510 exercises with European or desert terrain databases. Each exercise can contain as many as ten targets. (U-COFT Utilization Handbook, 1985).

Test Construction. The U-COFT Gunner's Test developed for this research consisted of eight shortened exercises from the U-COFT's TC/Gunner's training matrix, with each shortened exercise containing four target engagements. One target was friendly (an M2), making a total of 31 target engagements. The exercises were selected using a matrix sampling approach to match target conditions found in Table VIII of the M1 tank combat tables (FM 17-12-1). Table 1 shows the engagement conditions for Table VIII, while Table 2 lists the U-COFT exercises selected for the U-COFT Gunner's test.

Table VIII and the U-COFT test are similar in that half of the engagements in each are own vehicle stationary (defensive) vs. moving (offensive), and the two require an equal number of day channel and thermal (TIS) optics. Half of the Table VIII engagements are likewise short range (<1300m), as are half of the U-COFT targets (<1500m). The tests differ in that the U-COFT test is comprised only of main gun engagements fired by the gunner and contains no NBC conditions. The U-COFT test also includes a greater proportion of degraded conditions and makes use of the simulated battlefield distractions, e.g. friendly and enemy fire.

Dependent Measures. Nine performance measures were obtained from each engagement. These included Hit Rate which was defined as the number of hits divided by the number of targets presented. Other measures included First Round Hit Rate, Azimuth and Elevation errors, Target Identification (ID) time, which was the time from when the target appeared until the gunner said "identified," and Opening Time, which was the time from target appearance until the first round was fired.

The U-COFT software package reports three composite performance scores which were also recorded. Each of the scores is reported as a letter grade, A, B, C, or F with corresponding numerical values of 4.0, 3.0, 2.0, and 1.0. They are: 1) the Target Acquisition score which measures "skills required for the crew to accurately detect, identify, and classify targets," and is determined by target acquisition time and identification/classification errors (U-COFT Utilization Handbook, 1985); 2) the Reticle Aim score which assesses, "those skills required to lay the reticle on the proper aiming point, fire at, and destroy a given target(s)" and is computed from opening time, time to kill, and reticle aim error; and 3) the System Management score which measures "the ability to operate as a crew, utilizing the correct principles and techniques of gunnery," and counts pre-firing switch errors, ammunition errors, and excessive own vehicle exposure times.

Table 1.

Engagement conditions in Table VIII (FM 17-12-1)

	Crew member firing	Own vehicle	Target Number	Target Kind	Target Range	Vis	Optics	Weapon	Engage mode	Fire control malf	NBC
1)	Gunner	Stat	Mult	Stat Moving	<1300m <1300m	Day	GPS/Day	Main Gun Main Gun	Battle-sight	Comp, LRF	No
2)	Gunner TC	Stat	Simul	Stat Stat	<1100m <600m	Day	GPS/Day CWSS	Main Gun Cal .50	Prec	None	No
3)	Gunner	Moving	Mult	Troops	<600m <900m	Day	GPS/Day	Coax	Area	None	No
4)	Gunner	Moving	Mult	Stat Stat	>1400m >1400m	Day	GPS/Day	Main Gun	Prec	None	Yes
5)	Gunner	Moving	Mult	Moving Moving	>1400m >1400m	Day	GPS/Day	Main Gun	Prec	None	No
5A) Alt	Gunner	Moving	Mult	Stat Moving	>1400m >1400m	Day	GPS/Day	Main Gun	Prec	None	No
6)	TC	Stat	Single	Stat	>1400m	Night	GPSE/TIS	Main Gun	Prec	None	No
6A) Alt	Gunner	Stat	Single	Moving	>1700m	Night	GPS/TIS	Main Gun	Prec	None	No
7)	Gunner	Stat	Mult	Stat Stat	<1400m <1400m	Night	GPS/TIS	Main Gun	Prec	None	No
8)	Gunner	Moving	Mult	Stat Stat	<600m <600m	Night	GPS/TIS	Main Gun Coax	Prec	None	Yes
9)	Gunner	Moving	Mult	Stat Moving	<1500m <1500m	Night	GPS/TIS	Main Gun	Prec	None	No
10)	Gunner	Stat	Single	Stat	<1400m	Night with Illumination	GPS/Day	Main Gun	Prec	TIS	No

Note: Stationary (Stat), Multiple (Mult), Gunner's Primary Sight (GPS), Commander's Weapon Station Sight (CWSS), Gunner's Primary Sight Extension (GPSE), Thermal Imaging System (TIS), Precision (Prec), Computer System (Comp), Laser Rangefinder (LRF).

Procedure

The U-COFT Gunner's test was administered at the completion of other research evaluating training transfer between the U-COFT and MK-1 videodisc Gunnery Simulator (Witmer, in preparation). The gunners were randomly paired with the TCs and received 1 1/2 hours of U-COFT training during the MK1/U-COFT transfer study.

The U-COFT gunner's test began with eight practice engagements. The last four required use of the gunner's auxiliary sight (GAS) as there was simulated failure of the laser rangefinder (LRF), stabilization system, gunner's primary sight and computer system. The TC trained the gunner on use of the GAS and how to fire with manual lead and elevation.

Table 2.

Engagement conditions of U-COFT exercises used in this research

	Crew member firing	Own vehicle	Number	Target Kind	Target Range	Vis	Optics	Weapon	Engage mode	Fire control malf	NBC
1)	Gunner	Stat	Single	Moving	<1500m	Day	GPS/Day	Main Gun	Prec	None	No
2)	Gunner	Stat	Single	Stat	>1500m	Night	GPS/TIS	Main Gun	Prec	None	No
3)	Gunner	Moving	Single	Stat	>1500m	Dusk	GPS/Day	Main Gun	Prec	None	No
4)	Gunner	Moving	Mult	Moving Moving	<1500m <1500m	Night	GPS/TIS	Main Gun	Prec	None	No
5)	Gunner	Moving	Mult	Stat Stat	>1500m >1500m	Day	GPS/Day	Main Gun	Prec	None	No
6)	Gunner	Stat	Mult	Moving Moving	<1500m <1500m	Day	GAS/Day	Main Gun	Battle-sight	Comp, LRF Stabil & GPS	No
7)	Gunner	Stat	Mult	Stat Moving	>1500m >1500m	Night	GPS/TIS	Main Gun	Prec	None	No
8)	Gunner	Moving	Mult	Stat Stat	<1500m <1500m	Day-Fog	GPS/TIS	Main Gun	Battle-sight	LRF	No

Note: Stationary (Stat), Multiple (Mult), Gunner's Primary Sight (GPS), Gunner's Auxillary Sight (GAS), Thermal Imaging System (TIS), Precision (Prec), Computer System (Comp), Laser Rangefinder (LRF), Stabilization System (Stabil).

The eight U-COFT exercises with four engagements were sequentially presented with a short pause between each. During this time, the U-COFT I/O had to terminate the standard 10 engagement exercise, dump the printouts, and enter the six-digit code for the next subtest. This procedure was awkward and a few printouts were missed. Following a ten minute break, a retest was presented, which consisted of a different exercise order. The test and retest each took approximately 45 minutes, with the entire procedure lasting about two hours.

An attempt was made to minimize the effects of differential TC performance by having the I/O talk the TC onto the target. The I/O might have said, for example, "next target, a T-72, is left in 10 seconds." This modification theoretically reduced the variability of target identification times across TCs and minimized fire command errors. The standard U-COFT procedure also requires the gunner on defensive engagements to move his head out of the GPS, check the GAS to see that the gun has cleared the berm, say "driver stop", and then go back to the GPS. This procedure was omitted.

RESULTS AND DISCUSSION

Test Reliability

Mean scores for the tests and retests for each of the nine U-COFT measures are shown in Table 3, as is the Pearson r reliability coefficient. To the extent that the correlation coefficient approaches +1.0, a test is said to be reliable.

The reliability for six of the U-COFT measures was greater than .70, and for three of those measures, at least .80. Those reliabilities are encouraging and compare well with estimates of reliability of other tank gunnery measures (Powers, McCluskey, Haggard, Boycan, and Steinheiser, 1975; Eaton and Whalen, 1980). Reliability is important as it is a necessary component of valid tests. While high reliability does not guarantee validity, low reliability does guarantee low validity. Theoretically, the validity of a test cannot exceed the square of its reliability. If a test has a reliability of .50, for example, the validity can be no greater than .25. At best then, only 25% of this hypothetical test could be measuring "true" ability while 75% of score would be comprised of error.

Table 3

Mean test and retest scores, and reliability (Pearson r) for the nine U-COFT measures

Variable	Test Mean	Retest Mean	r
Target ID Time (sec)	5.50	4.70	.87
Reticle Aim Score (1-4)	1.92	2.03	.83
Hit Rate	.67	.69	.80
Target Acquisition Score (1-4)	2.65	2.83	.76
First Round Hit Rate	.59	.60	.72
Opening Time (sec)	16.10	15.20	.72
Azimuth Error (mils)	1.16	1.14	.42
System Management Score (1-4)	2.58	2.57	.11
Elevation Error (mils)	.53	.62	-.07

What has been said about reliability and validity thus far is the view from traditional psychometry. Alternative approaches to reliability have developed "consistency" measures which partition sources of error as being attributable to different test administrators or administrations, differences in scoring, differences in the way subjects react to the test, or various combinations of these factors (Cattell, 1964). Device-mediated tests such as with the U-COFT minimize several of these sources by standardizing test

administration and scoring procedures. Attempts are likewise being made to increase the reliability of live-fire gunnery exercises at Grafenwehr Range 20 and elsewhere by computerizing targets and procedures.

Table 3 shows poor reliability for Azimuth errors, Elevation errors and the U-COFT System Management scores. The unreliability of the elevation and azimuth error appears to have resulted from some extreme scores. Although mean elevation error of .57 mils is within the hit range of many targets, some individual U-COFT exercise elevation errors exceeded 7.0. These extreme scores are not likely be found with more experienced gunners. Reticle Aim error is potentially the most sensitive and useful measure of gunnery performance, and should therefore be reexamined with other more experienced sample of gunners.

The unreliability of System Management errors probably resulted from a part of the U-COFT procedure which differed from "real world" gunnery. During the exercises, the gunners were instructed to leave the gun select switch on MAIN GUN. On defensive engagements when a crew was "killed" as the result of being exposed too long, the gun select switch automatically reset to SAFE. Switching errors may have resulted from failure to put the switch back on MAIN GUN. Other system management errors resulted from incorrect ammo select switch settings.

The feasibility of using U-COFT as a testing device depends, in part, on its ease of the administration and scoring. The three U-COFT scores, Reticle Aim, Target Acquisition, and System Management which are printed after each exercise could well be used as performance measures. The composite mean of these scores yielded a test-retest reliability of .82, and correlated .87 with Hit Rate. These data substantiate the internal validity of the U-COFT scoring procedure.

One minor difficulty with the U-COFT for training and testing is that it includes dispersion rounds. A gunner may have a perfect sight picture, fire within the required time limits, and still miss the target. The opposite also occurred, although less frequently. While the dispersion rounds were likely included to match live ammunition characteristics, they reduce the reliability of U-COFT as a testing medium by randomly altering a soldier's score. In addition, the dispersion rounds provide random bad feedback during training, which is likely to be most detrimental early in training. Software updates should include the capability of turning off the dispersion rounds.

Test Utility

Despite attempts to minimize TC effects, U-COFT performance differed as a function of TCs, i.e., COFT-experienced, Sr NCO, and Jr NCO. Gunners paired with the COFT-experienced TC had a combined test and retest Hit Rate of 74%, while the gunners paired with the Sr NCO shot 64%, and those with the Jr NCO shot 63%. An analysis of variance showed those differences to be statistically significant, $F(2,28) = 3.89$, $p < .05$. Similar significant differences as a function of TCs also were found for First Round Hit Rate, ID time, and the Reticle Aim, Target Acquisition, and System Management scores.

Hit Rate data were recomputed to reflect these changes over sessions. Figure 1 shows changes in Hit Rates for the three TCs over the duration of the experiment. Each TC session reflects the mean performance of three different gunners. Figure 1 shows that Hit Rate was heavily influenced by TC performance. The first three gunners paired with the Jr NCO hit 48% of the targets while his last three gunners hit 77%. Each testing session gave the TC about 3 1/2 hours of U-COFT experience: two hours in this experiment and 1 1/2 hours in the immediately preceeding experiment. Three TC sessions, therefore, corresponded to about 10 hours of U-COFT time. The horizontal axis in Figure 1 could alternatively be labeled 10, 20, 30 and 40 hours of TC COFT time. These data show that the COFT-experienced TC's contribution to hit performance reached asymptote at between 10 and 20 hours, while the Jr NCO's contribution had not stabilized at 40 hours.

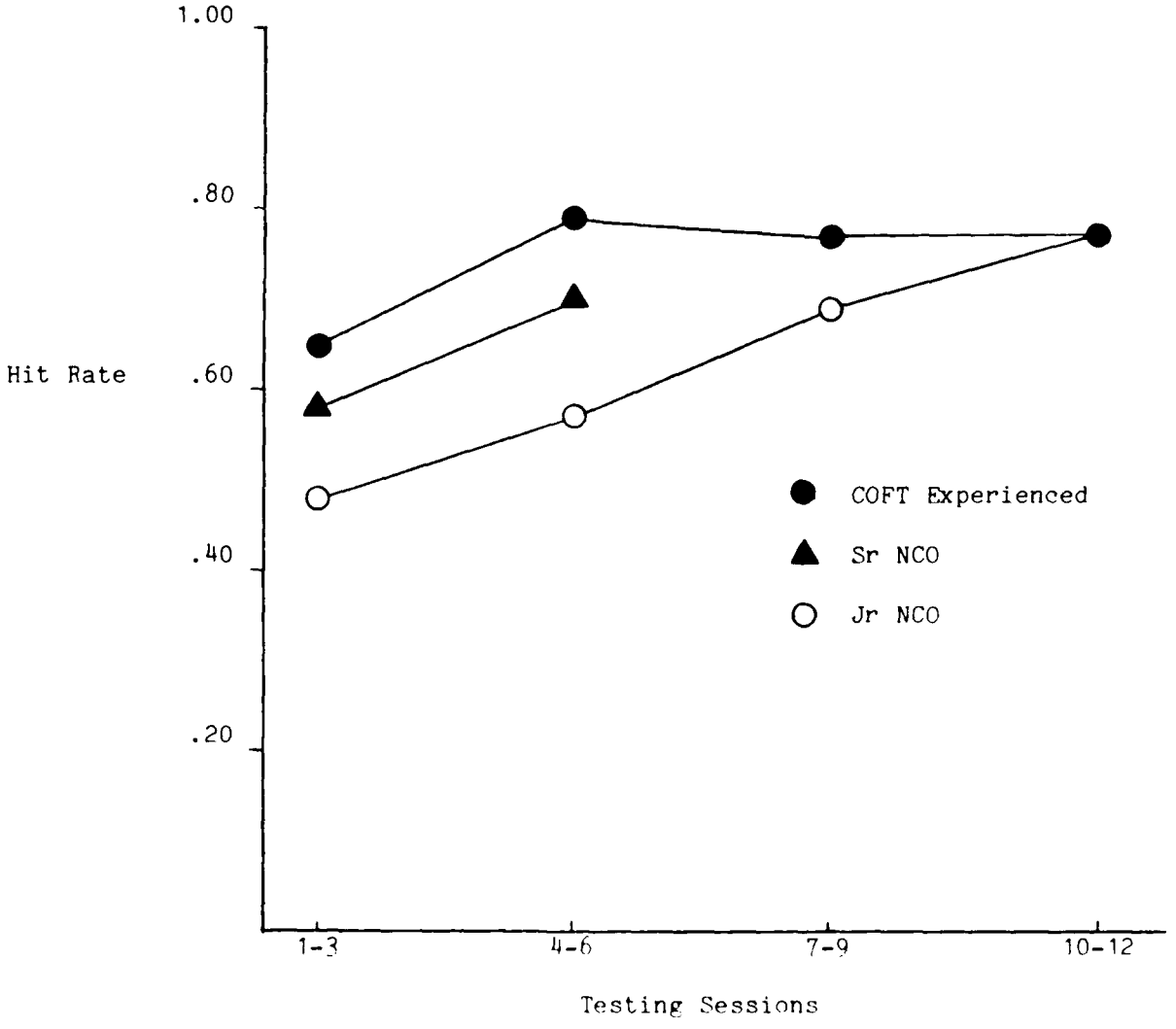

Figure 1. Hit Rate as a Function of Testing Sessions and TCs

Two factors may have contributed to this pattern. When target ID time is plotted as a function of TC sessions, improvement in ID time for the three TCs essentially mirrors Hit Rate performance. This finding suggests that faster ID times lead to higher Hit Rates, although a regression analysis discussed later suggests otherwise. The second factor involves changes in the gunners' reticle aim errors over TC sessions, as improvements in reticle aim also showed the same general pattern as Hit Rate and ID times. Improved reticle aim over TC sessions is believed to have resulted from the TCs improving as trainers of the novice gunners. The TCs learned what errors were typically being made by the gunners, and better trained the later gunners in these areas, e.g., tracking and use of the gunner's auxillery sight.

Improvement as a trainer was particularly noticeable for the Jr NCO. He nervously barked his fire commands in the earlier sessions, and paid little attention to the gunner. In later sessions, he readily assisted the I/O in the U-COFT orientation and praised the gunners during the breaks. General Electric officials observed a similar phenomenon during the U-COFT validation tests. TCs and gunners at first blamed each other for misses, but became increasingly cooperative as the training progressed.

The results summarized in Figure 1 suggest that:

a. Stabilization of TC performance on the U-COFT takes a long time. The performance of the Jr NCO, for example, would have stabilized only after about 40 hours, which is enough time to have progressed about halfway through the U-COFT training matrix.

b. The most efficient U-COFT training strategy might train TCs alone, before TC/gunner pairs. U-COFT system familiarity and reduced TC target ID time should precede training with the gunner. One solution is to give the TC institutional training on the Institutional-Conduct of Fire Trainer (I-COFT) as part of the Armor Basic Non-commissioned Officers Course.

c. Gunner's test scores on the U-COFT result from a number of factors other than the gunners' skills in laying and firing the gun. The extent to which improved scores result from gunnery skill and familiarity with the U-COFT in unknown.

d. U-COFT test scores are highly influenced by the TCs' ability to train their crews. The U-COFT could better be used to train the trainer to train if, for example, TC/Gunner interactions were noted by the I/O and discussed with the crew as part of the exercise debriefing.

e. The U-COFT is not appropriate as a medium for assessing the proficiency of gunners alone if TC performance has not stabilized. Units might accomplish this by having one Master Gunner serve as TC for testing all gunners. The U-COFT is better suited for testing the proficiency of TC/gunner pairs than for testing gunners.

The utility of a test is closely related to its cost-effectiveness. A test is cost-effective when improved decisions or financial savings resulting from the test exceed the cost of developing and administering the test. Tests

which result in ceiling effects, i.e., many scores are near perfect, provide little information for making decisions. It is important therefore in test construction to include a majority of items which are neither too easy nor difficult.

Table 4 lists the U-COFT exercises used in this research along with their combined test and retest means and standard deviations. Hit Rate performance for U-COFT Exercises 1-3 approached ceiling levels eventhough the experimental gunners were M60A3 loaders and drivers with little gunners' experience. A ceiling effect would be expected if the easier target engagements were included on a test with experienced TC/gunner pairs, with the restricted range of scores contributing to lower test reliability. The high Hit Rate for the long range single stationary targets does, however, demonstrate the relative ease of using the M1 fire control system when it is fully operational.

The utility of the U-COFT as a testing medium is further enhanced by the fact that test conditions are not restricted by range, equipment and ammunition limitations. An outstanding feature of the U-COFT is its ability to train and test under various degraded conditions. Performance on U-COFT Exercise Number 6, in which there was simulated failure of the computerized fire control system, is seen in Table 4 to be considerably below the other conditions. Future U-COFT tests might place even a greater emphasis on evaluating degraded gunnery performance, e.g., manual control conditions.

Model Gunnery Performance

A possible use of U-COFT test scores in addition to assessing proficiency is in the development and refinement of combat attrition models. The value of these complex models has been limited in the past by poor parameter estimates. The estimates have largely come from battle data which are sketchy at best, as armies engaged in war are interested in a number of factors in addition to gathering quality data for combat models. The U-COFT has the capability of gathering reliable estimates under a number of conditions, including future battlefield conditions. As such, the U-COFT can be used to evaluate alternative future training strategies.

Stepwise regression analyses were performed in the present research to help understand what factors were underlying the various performance measures. Table 5 shows the results of several analyses. The Hit Rate measure was found to be primarily determined by reticle aim accuracy, as target ID time and Opening time did not significantly contribute to the prediction of Hit Rate (Equation 1). In the previous discussion of possible factors affecting changes in Hit Rate over TC sessions, changes in both target ID time and reticle aim accuracy were shown to parallel changes in Hit Rate. The regression analysis suggests, however, that changes in target ID time are unrelated to changes in Hit Rate, once the effects of aiming error are taken into account. This finding supports the explanation that improvement in Hit Rate over TC sessions was primarily due to improved reticle aim, and that this is attributed to better training of gunners over TC sessions.

Table 4

U-COFT exercises with hit rate means and standard deviations

U-COFT Exercise #		Description	Mean	Standard Deviation
1)	313110	Stationary tank - short range single moving targets (Gunner - Precision - GPS - Normal - Day)	.83	.15
2)	322210	Stationary tank - long range single stationary targets (Gunner - Precision - GPS - Normal - Night)	.91	.20
3)	324120	Moving tank - long range single stationary targets (Gunner - Precision - GPS - Normal - Dusk)	.87	.16
4)	335230	Moving tank - short range multiple moving targets (Gunner - Precision - GPS - Normal - Night - Clutter - Friendly & Enemy Fire - Friendly Targets)	.59	.20
5)	344130	Moving tank - long range multiple stationary targets (Gunner - Precision - GPS - Normal - Day - Friendly & Enemy Fire)	.57	.18
6)	333610	Stationary tank - short range multiple moving targets (Gunner - Battlesight - GAS - Emergency - Day - Malf LRF - STAB - GPS - Comp)	.18	.20
7)	346210	Stationary tank - long range moving & stationary targets (Gunner - Precision - GPS - Normal Night)	.73	.24
8)	334520	Moving tank - short range multiple stationary targets (Gunner - Battlesight - GPS - Normal - Day Fog - Malf: LRF)	.73	.23

The regression analyses also suggest that Opening time is the result of target ID time and not the result of Reticle Aim Accuracy (Equation 2). ID time is primarily a measure of the TC's ability to acquire the target and to make a gross lay of the gun. The U-COFT Reticle Aim score while computed from Opening Time, Time to Kill, and Reticle Aim Error is shown to be primarily related to Hit Rate and to a lesser extent ID time (Equation 3). The point here is that the scale properties of the U-COFT composite measures are not easily discernible, even when one knows the algorithms listed in the U-COFT Utilization Handbook. This knowledge is, however, important. Future research with the U-COFT will require a complete understanding of how it measures various parameters, processes them, and reports them.

Table 5

Results of stepwise regression analyses

Criterion		Predictors				R	R^2
(1) Hit Rate		Azimuth Error	Elevation Error	ID Time	Opening Time		
	Beta	-.52	-.49	-	-	.79	.62
(2) Opening Time		Azimuth Error	Elevation Error	ID Time	Hits		
	Beta	-	-	.79	-	.79	.62
(3) Reticle Aim Scores		Hits	Azimuth Error	Elevation Error	ID Time	Opening Time	
	Beta	.72	-	-	.31	-	.94	.88

Note: Includes only those variables contributing significantly to multiple R.

CONCLUSIONS

The U-COFT test used in this research was reliable, and is therefore a potentially valid means for assessing gunnery proficiency. Test-retest reliability coefficients exceeding .80 were found for a number of performance measures including Hit Rate and target ID time, and these values are well within the acceptable psychometric range for military tests. In addition to assessing proficiency, U-COFT tests can be used in the development of combat attrition models and as criteria against which other less expensive part-task tests can be validated.

Device-mediated tests such as with the U-COFT offer certain advantages over hands-on performance tests including standardized administration and scoring procedures, and the inexpensive capability of building longer, more varied tests. The U-COFT tests can also be used to partial out the relative contribution of individual crew members in whole-task gunnery engagements. If the U-COFT is to be used to assess only gunners' performance, however, TCs must be trained to standardized levels of performance.

The U-COFT proficiency test developed for this research mirrored target conditions in the M1 tank Table VIII, but this may not be necessary. As Boldovici (1979) noted, the lowest level of "enabling skills" required for

gunnery performance, e.g., psychomotor skills and system procedural knowledge, are highly redundant across engagement conditions. As a result, it is unlikely that a crew who is relatively good at long range moving targets at night would be poor at short range stationary targets in daylight. Future U-COFT tests should select exercises which avoid ceiling effects for the population of soldiers to be evaluated, and depending on the testing purpose, might include more exercises in degraded modes.

Improved evaluation of tank gunnery skills with the U-COFT can lead to a stronger Armor force. Skills and abilities other than those measured by the U-COFT are, however, equally important and should not be overlooked. Graham and Black (1985) found that soft skills, e.g., ability to train, ability to communicate, and general leadership qualities, were predominately identified as the distinguishing characteristics of TC excellence. Likewise, the results described here showed the TC's ability to train his gunner heavily influenced gunnery performance. Continuing development of evaluation batteries which assess both hard combat-oriented skills and soft leadership skills is needed, if the force is truly to become an Army of excellence.

References

Boldovici, J.A. (1979). Analyzing Tank Gunnery Engagements for Simulator-Based Process Measurements. Army Research Institute Research Report 1227.

Boldovici, J.A. & Sabat, S.R. (1985). Measuring Transfer from Training Devices to Weapon Systems. Paper presented at NATO Symposium on the Transfer of Training to Operational Military Systems, Brussels, Belgium.

Cattell, R.B. (1964). Validity and Reliability: A Proposed More Basic Set of Concepts. Journal of Educational Psychology, 55, 1-22.

Eaton, N.K. & Whalen, D.E. (1980). Accuracy of Tank Gunnery Scoring. Army Research Institute Working Paper FKFU 80-4.

Graham, S.E. & Black, B.A. (1985). Defining and Assessing Tank Commander Excellence. Army Research Institute Research Report 1401. (AD A156 804)

Powers, R.P., McCluskey, M.R., Haggard, D.F., Boycan G.G., & Steinheiser, F. (1975). Determination of the Contribution of Live Firing to Weapons Proficiency. Human Resources Research Organization Report FR-CD(C). 75-1.

Unit-Conduct of Fire Trainer Instructor's Utilization Handbook (1985). (Daytona Beach, FL.) General Electric Company.

Witmer, B.G. (in Prep). Unit-Conduct of Fire Trainer (U-COFT) and Videodisc Gunnery Simulator (VIGS) Cross-Training and Training Transfer. Army Research Institute-Fort Knox Field Unit Working Paper.